Week 16

Shaun Stearns

9/26/2019

Multiple Regression 1

Lowest Weigh-in

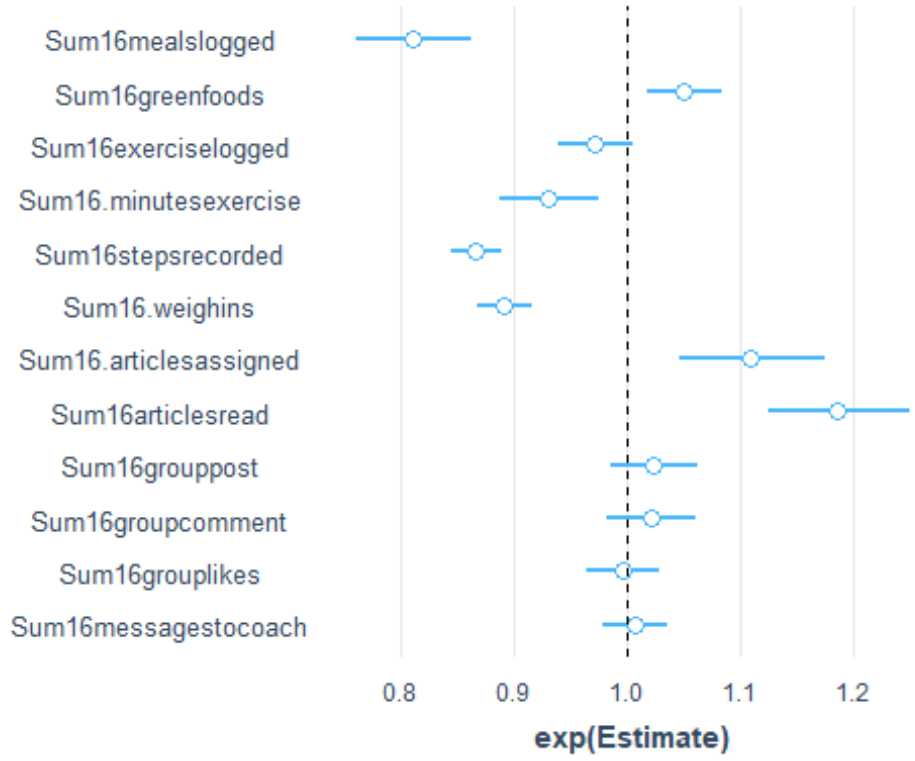
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(10,18,21,24,27,30,33,36,39,42,45,48,51)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(Lowestweighinkg ~., DF.t)
```

MODEL INFO:
Observations: 7138
Dependent variable: Lowestweighinkg
Type: OLS linear regression

MODEL FIT:
$F(12,7125) = 33.16$, $p = 0.00$
$R^2 = 0.05$
Adj. $R^2 = 0.05$

Standard errors: OLS

	Est.	2.5%	97.5%	t val.	p	partial.r	part.r
(Intercept)	0.00	-0.02	0.02	0.00	1.00		
Sum16mealslogged	-0.21	-0.27	-0.15	-6.66	0.00	-0.08	-0.08
Sum16greenfoods	0.05	0.02	0.08	3.33	0.00	0.04	0.04
Sum16exerciselogged	-0.03	-0.06	0.00	-1.69	0.09	-0.02	-0.02
Sum16.minutesexercise	-0.07	-0.10	-0.04	-4.68	0.00	-0.06	-0.05
Sum16stepsrecorded	-0.14	-0.17	-0.12	-10.03	0.00	-0.12	-0.12
Sum16.weighins	-0.12	-0.14	-0.09	-8.24	0.00	-0.10	-0.09
Sum16.articlesassigned	0.10	0.05	0.16	3.62	0.00	0.04	0.04
Sum16articlesread	0.17	0.12	0.22	6.21	0.00	0.07	0.07
Sum16grouppost	0.02	-0.01	0.06	1.39	0.16	0.02	0.02
Sum16groupcomment	0.02	-0.02	0.06	1.07	0.29	0.01	0.01
Sum16grouplikes	-0.00	-0.04	0.03	-0.25	0.80	-0.00	-0.00
Sum16messagestocoach	0.01	-0.02	0.04	0.42	0.68	0.00	0.00



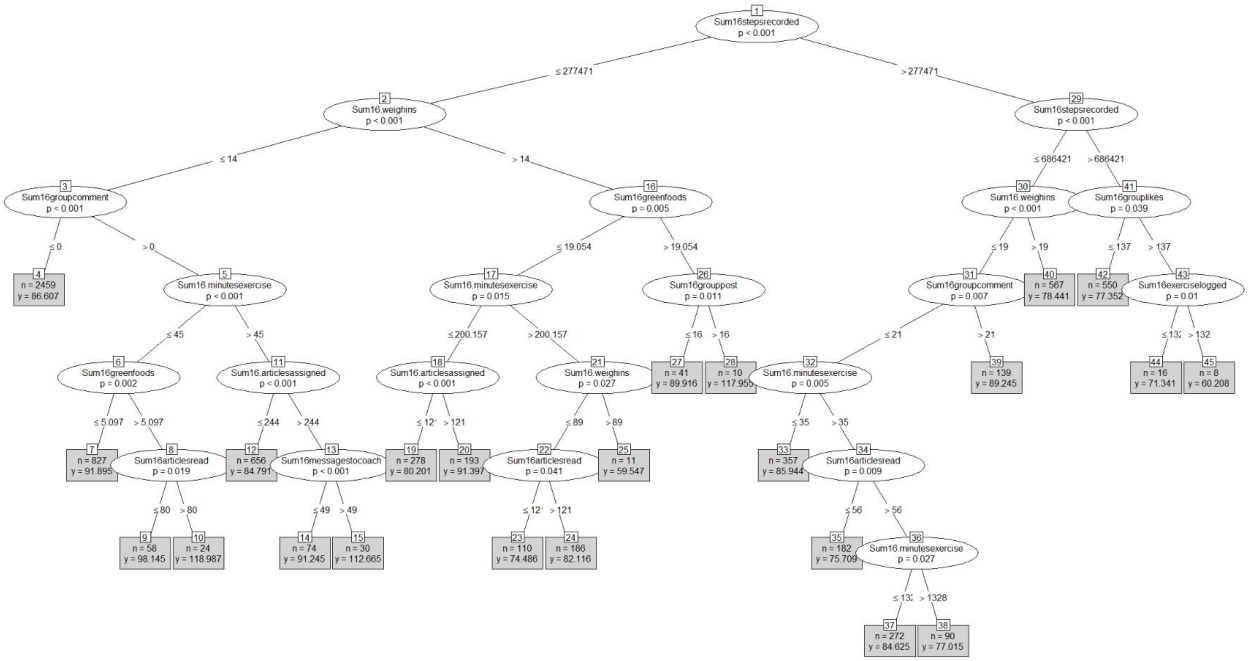
Regression Tree 1

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(10,18,21,24,27,30,33,36,39,42,45,48,51)]
model <- train(
  Lowestweighinkg ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results
```

```
##   mincriterion     RMSE   Rsquared      MAE   RMSESD RsquaredSD     MAESD
## 1         0.95 20.17682 0.04307264 15.61753 0.718905 0.02035198 0.4633603
```

Tree Model

```
plot(model$finalModel, type = "simple")
```



User Engagement and Lowest Weigh-in based on 23 Terminal Nodes (Left to Right)

1	Lowest Weigh-in avg 86.61 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins ≤ 14, WK 16 Group Comments ≤ 0**
2	Lowest Weigh-in avg 91.90 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins ≤ 14, WK 16 Group Comments > 0, WK 16 Min Exercise ≤ 45, WK 16 Green Foods ≤ 5.10**
3	Lowest Weigh-in avg 98.15 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins ≤ 14, WK 16 Group Comments > 0, WK 16 Min Exercise ≤ 45, WK 16 Green Foods > 5.10, WK 16 Articles Read ≤ 80**
4 High	Lowest Weigh-in avg 118.99 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins ≤ 14, WK 16 Group Comments > 0, WK 16 Min Exercise ≤ 45, WK 16 Green Foods > 5.10, WK 16 Articles Read > 80**
5	Lowest Weigh-in avg 84.79 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins ≤ 14, WK 16 Group Comments > 0, WK 16 Min Exercise > 45, WK 16 Articles Assigned ≤ 244**
6	Lowest Weigh-in avg 91.25 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins ≤ 14, WK 16 Group Comments > 0, WK 16 Min Exercise > 45, WK 16 Articles Assigned > 244, WK 16 Messages to Coach ≤ 49**
7 2nd High	Lowest Weigh-in avg 112.67 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins ≤ 14, WK 16 Group Comments > 0, WK 16 Min Exercise > 45, WK 16 Articles Assigned > 244, WK 16 Messages to Coach > 49**
8	Lowest Weigh-in avg 80.20 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins > 14, WK 16 Green Foods ≤ 19.05, WK 16 Min Exercise ≤ 200.16, WK 16 Articles Assigned ≤ 121**
9	Lowest Weigh-in avg 91.40 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins > 14, WK 16 Green Foods ≤ 19.05, WK 16 Min Exercise ≤ 200.16, WK 16 Articles Assigned > 121**
10	Lowest Weigh-in avg 74.49 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins > 14, WK 16 Green Foods ≤ 19.05, WK 16 Min Exercise > 200.16, WK 16 Weigh-ins ≤ 89, WK 16 Articles Read ≤ 121**
11	Lowest Weigh-in avg 82.12 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins > 14, WK 16 Green Foods ≤ 19.05, WK 16 Min Exercise > 200.16, WK 16 Weigh-ins ≤ 89, WK 16 Articles Read > 121**
12 Low	Lowest Weigh-in avg 59.58 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins > 14, WK 16 Green Foods ≤ 19.05, WK 16 Min Exercise > 200.16, WK 16 Weigh-ins > 89**

13	Lowest Weigh-in avg 89.92 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins > 14, WK 16 Green Foods > 19.05, WK 16 Group Posts ≤ 16**
14	Lowest Weigh-in avg 117.96 kg = **WK 16 Steps Recorded ≤ 277471, WK 16 Weigh-ins > 14, WK 16 Green Foods > 19.05, WK 16 Group Posts > 16**
15	Lowest Weigh-in avg 85.94 kg = **WK 16 Steps Recorded > 277471, WK 16 Steps Recorded ≤ 686421, WK 16 Weigh-ins ≤ 19, WK 16 Group Comments ≤ 21, WK 16 Min Exercise ≤ 35**
16	Lowest Weigh-in avg 75.71 kg = **WK 16 Steps Recorded > 277471, WK 16 Steps Recorded ≤ 686421, WK 16 Weigh-ins ≤ 19, WK 16 Group Comments ≤ 21, WK 16 Min Exercise > 35, WK 16 Articles Read ≤ 56**
17	Lowest Weigh-in avg 84.63 kg = **WK 16 Steps Recorded > 277471, WK 16 Steps Recorded ≤ 686421, WK 16 Weigh-ins ≤ 19, WK 16 Group Comments ≤ 21, WK 16 Min Exercise > 35, WK 16 Articles Read > 56, WK 16 Min Exercise ≤ 1328**
18	Lowest Weigh-in avg 77.02 kg = **WK 16 Steps Recorded > 277471, WK 16 Steps Recorded ≤ 686421, WK 16 Weigh-ins ≤ 19, WK 16 Group Comments ≤ 21, WK 16 Min Exercise > 35, WK 16 Articles Read > 56, WK 16 Min Exercise > 1328**
19	Lowest Weigh-in avg 89.25 kg = **WK 16 Steps Recorded > 277471, WK 16 Steps Recorded ≤ 686421, WK 16 Weigh-ins ≤ 19, WK 16 Group Comments > 21**
20	Lowest Weigh-in avg 78.44 kg = **WK 16 Steps Recorded > 277471, WK 16 Steps Recorded ≤ 686421, WK 16 Weigh-ins > 19**
21	Lowest Weigh-in avg 77.35 kg = **WK 16 Steps Recorded > 277471, WK 16 Steps Recorded > 686421, WK 16 Group Likes ≤ 137**
22	Lowest Weigh-in avg 71.34 kg = **WK 16 Steps Recorded > 277471, WK 16 Steps Recorded > 686421, WK 16 Group Likes > 137, WK 16 Exercises Logged ≤ 132**
23 2nd Low	Lowest Weigh-in avg 60.21 kg = **WK 16 Steps Recorded > 277471, WK 16 Steps Recorded > 686421, WK 16 Group Likes > 137, WK 16 Exercises Logged > 132**

Random Forest 1

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(10,18,21,24,27,30,33,36,39,42,45,48,51)]
rf <- randomForest(Lowestweighinkg ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```

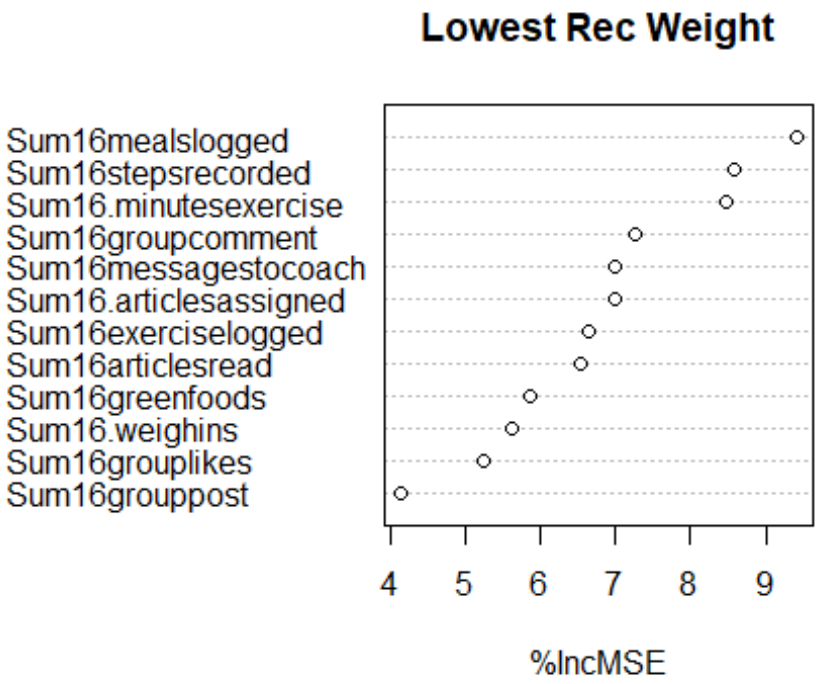
Signicance Testing

```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##   Model R-square:  -0.1457512
##   Random R-square:  -0.2671981
##   Random R-square variance:  0.0001450278
```

Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "Lowest Rec Weight")
```



Multiple Regression 2

Difference Between First Weigh-in and Lowest Weigh-in

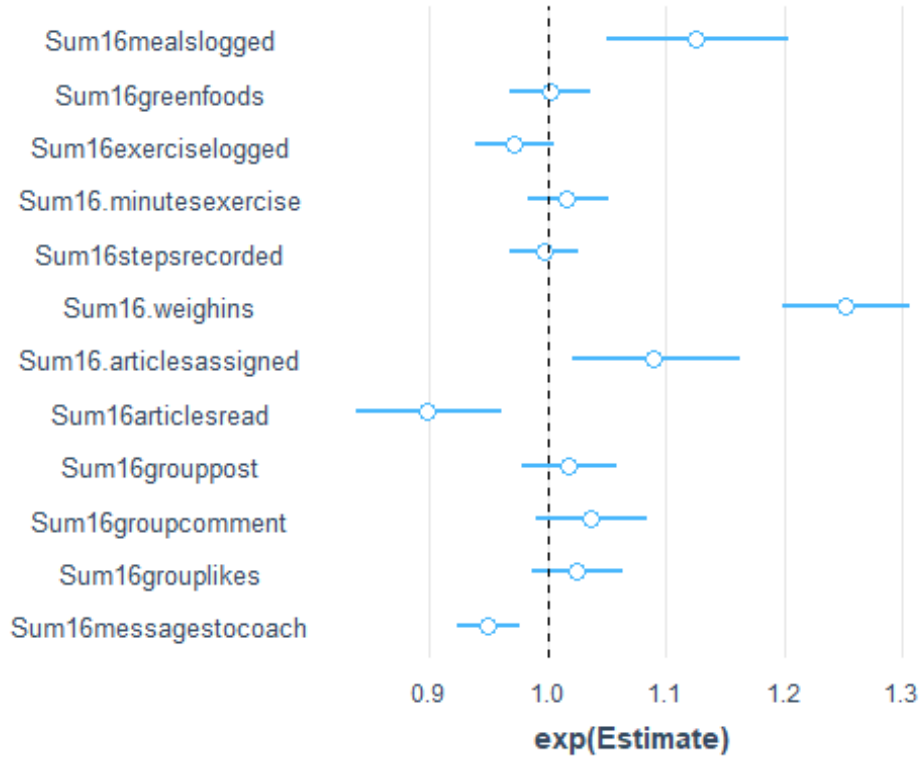
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(11,18,21,24,27,30,33,36,39,42,45,48,51)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(AbsDiffFirstWeighinkg ~., DF.t)
```

MODEL INFO:
Observations: 7138
Dependent Variable: AbsDiffFirstWeighinkg
Type: OLS linear regression

MODEL FIT:
$F(12,7125) = 57.44$, $p = 0.00$
$R^2 = 0.09$
Adj. $R^2 = 0.09$

Standard errors: OLS

	Est.	2.5%	97.5%	t val.	p	partial.r	part.r
(Intercept)	-0.00	-0.02	0.02	-0.00	1.00		
Sum16mealslogged	0.12	0.06	0.18	3.79	0.00	0.04	0.04
Sum16greenfoods	0.00	-0.03	0.03	0.12	0.91	0.00	0.00
Sum16exerciselogged	-0.03	-0.06	0.00	-1.74	0.08	-0.02	-0.02
Sum16.minutesexercise	0.02	-0.01	0.05	1.12	0.26	0.01	0.01
Sum16stepsrecorded	-0.00	-0.03	0.03	-0.17	0.86	-0.00	-0.00
Sum16.weighins	0.22	0.20	0.25	16.32	0.00	0.19	0.18
Sum16.articlesassigned	0.09	0.03	0.14	3.08	0.00	0.04	0.03
Sum16articlesread	-0.11	-0.16	-0.05	-3.95	0.00	-0.05	-0.04
Sum16grouppost	0.02	-0.01	0.05	1.07	0.29	0.01	0.01
Sum16groupcomment	0.04	-0.00	0.07	1.90	0.06	0.02	0.02
Sum16grouplikes	0.02	-0.01	0.06	1.45	0.15	0.02	0.02
Sum16messagestocoach	-0.05	-0.08	-0.02	-3.38	0.00	-0.04	-0.04



Regression Tree 2

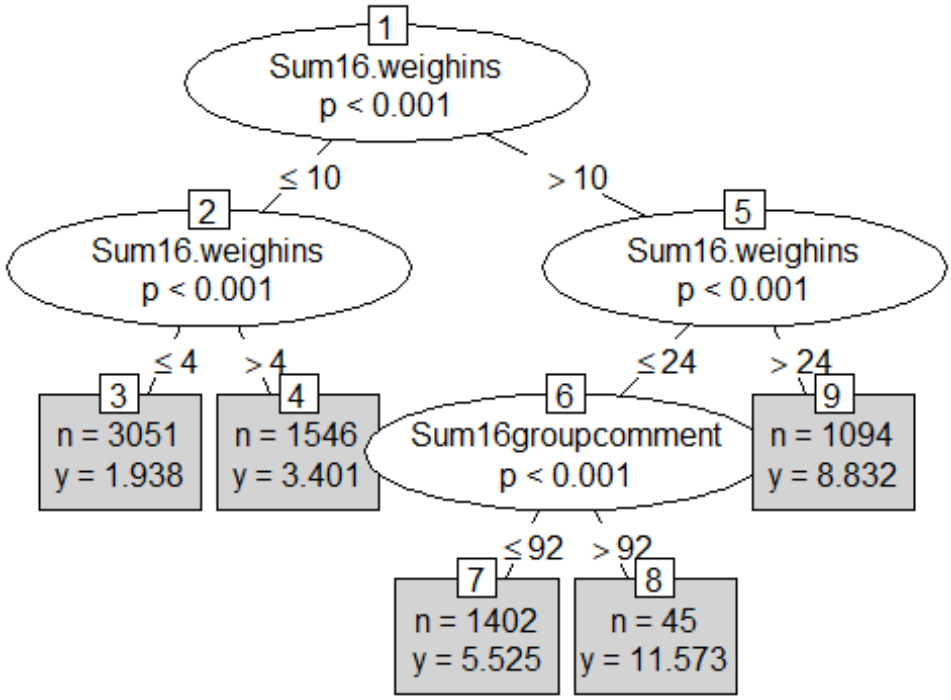
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header = TRUE)
```

```
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(11,18,21,24,27,30,33,36,39,42,45,48,51)]
model <- train(
  AbsDiffFirstWeighinkg ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results
```

```
##   mincriterion     RMSE   Rsquared      MAE    RMSESD RsquaredSD     MAESD
## 1         0.95 7.754058 0.09137554 3.862621 0.5712898  0.0245084 0.1510953
```

Tree Model

```
plot(model$finalModel, type = "simple")
```



User Engagement and Difference Between First Weigh-in and Lowest Weigh-in based on 5 Terminal Nodes (Left to Right)

1 High	Diff 1st Weight Lost avg 1.94 kg = **WK 16 Weigh-ins ≤ 4**
2	Diff 1st Weight Lost avg 3.40 kg = **WK 16 Weigh-ins > 4 & ≤ 10**
3	Diff 1st Weight Lost avg 5.53 kg = **WK 16 Weigh-ins > 10 & ≤ 24, WK 16 Group Comments ≤ 92**
4 Low	Diff 1st Weight Lost avg 11.57 kg = **WK 16 Weigh-ins > 10 & ≤ 24, WK 16 Group Comments > 92**
5	Diff 1st Weight Lost avg 8.83 kg = **WK 16 Weigh-ins > 24**

Random Forest 2

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(11,18,21,24,27,30,33,36,39,42,45,48,51)]
rf <- randomForest(AbsDiffFirstWeighinkg ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```

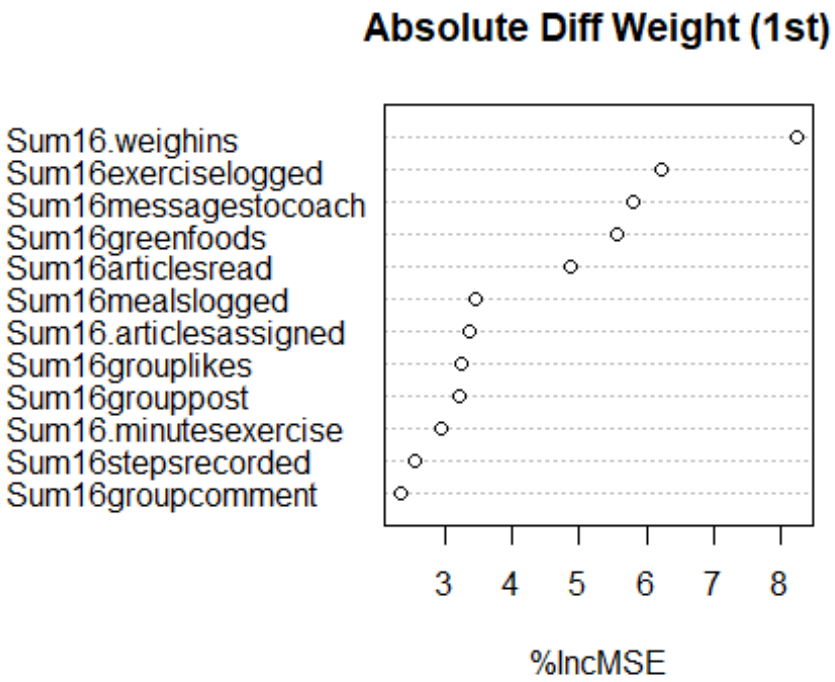
Signicance Testing

```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##   Model R-square:  -0.112512
##   Random R-square:  -0.2778766
##   Random R-square variance:  0.0005185674
```

Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "Absolute Diff Weight (1st)")
```



Multiple Regression 3

Differnece Between Initial Weigh-in and Lowest Weigh-in

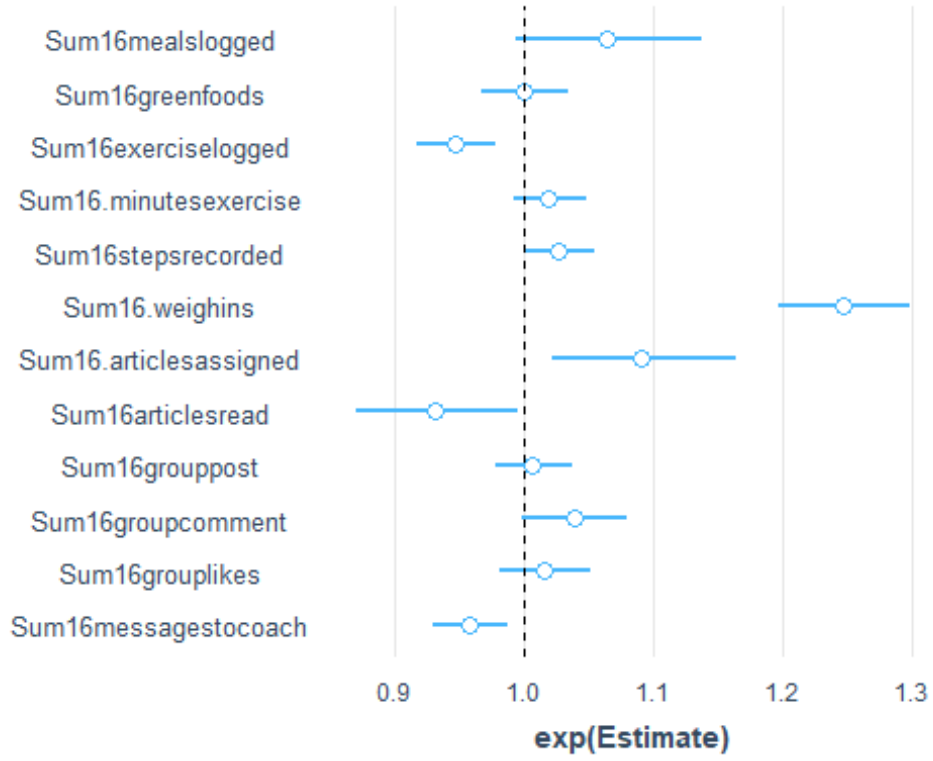
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(12,18,21,24,27,30,33,36,39,42,45,48,51)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(AbsDiffInitWeighinkg ~., DF.t)
```

```
MODEL INFO:
Observations: 7138
Dependent Variable: AbsDiffInitWeighinkg
Type: OLS linear regression

MODEL FIT:
F(12,7125) = 49.03, p = 0.00
R² = 0.08
Adj. R² = 0.07

Standard errors: OLS
```

	Est.	2.5%	97.5%	t val.	p	partial.r	part.r
(Intercept)	-0.00	-0.02	0.02	-0.00	1.00		
Sum16mealslogged	0.06	0.00	0.12	1.97	0.05	0.02	0.02
Sum16greenfoods	-0.00	-0.03	0.03	-0.01	0.99	-0.00	-0.00
Sum16exerciselogged	-0.05	-0.09	-0.02	-3.28	0.00	-0.04	-0.04
Sum16.minutesexercise	0.02	-0.01	0.05	1.28	0.20	0.02	0.01
Sum16stepsrecorded	0.03	-0.00	0.05	1.88	0.06	0.02	0.02
Sum16.weighins	0.22	0.19	0.25	15.94	0.00	0.19	0.18
Sum16.articlesassigned	0.09	0.03	0.14	3.09	0.00	0.04	0.04
Sum16articlesread	-0.07	-0.12	-0.02	-2.63	0.01	-0.03	-0.03
Sum16grouppost	0.01	-0.03	0.04	0.40	0.69	0.00	0.00
Sum16groupcomment	0.04	0.00	0.08	1.96	0.05	0.02	0.02
Sum16grouplikes	0.02	-0.02	0.05	0.93	0.35	0.01	0.01
Sum16messagestocoach	-0.04	-0.07	-0.01	-2.79	0.01	-0.03	-0.03



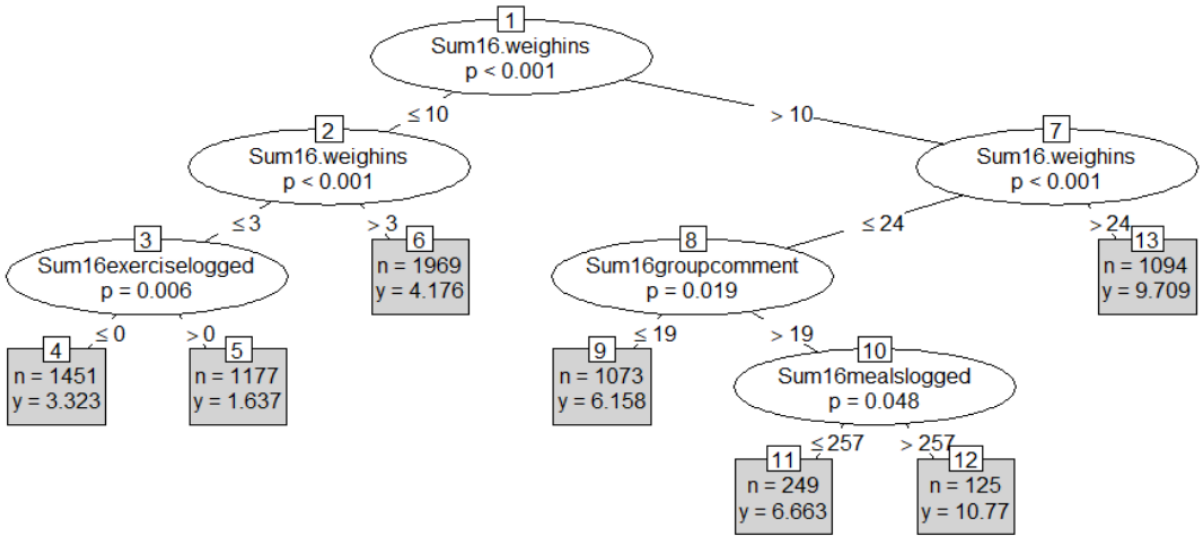
Regression Tree 3

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(12,18,21,24,27,30,33,36,39,42,45,48,51)]
model <- train(
  AbsDiffInitWeighinkg ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results
```

```
##   mincriterion     RMSE   Rsquared      MAE    RMSESD RsquaredSD     MAESD
## 1         0.95 8.398452 0.08560493 4.342875 0.8597068 0.02460985 0.2320555
```

Tree Model

```
plot(model$finalModel, type = "simple")
```



User Engagement and Difference Between Initial Weigh-in and Lowest Weigh-in based on 7 Terminal Nodes (Left to Right)

1	Diff Initial Weight Lost avg 3.32 kg = **WK 16 Weigh-ins ≤ 0**
2 Low	Diff Initial Weight Lost avg 1.64 kg = **WK 16 Weigh-ins > 0 & ≤ 3**
3	Diff Initial Weight Lost avg 4.18 kg = **WK 16 Weigh-ins > 3 & ≤ 10**
4	Diff Initial Weight Lost avg 6.16 kg = **WK 16 Weigh-ins > 10 & ≤ 24, WK 16 Group Comments ≤ 19**
5	Diff Initial Weight Lost avg 6.66 kg = **WK 16 Weigh-ins > 10 & ≤ 24, WK 16 Group Comments > 19, WK 16 Meals Logged ≤ 257**
6 High	Diff Initial Weight Lost avg 10.77 kg = **WK 16 Weigh-ins > 10 & ≤ 24, WK 16 Group Comments > 19, WK 16 Meals Logged > 257**
7	Diff Initial Weight Lost avg 9.71 kg = **WK 16 Weigh-ins > 24**

Random Forest 3

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(12,18,21,24,27,30,33,36,39,42,45,48,51)]
rf <- randomForest(AbsDiffInitWeighinkg ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```

Signicance Testing

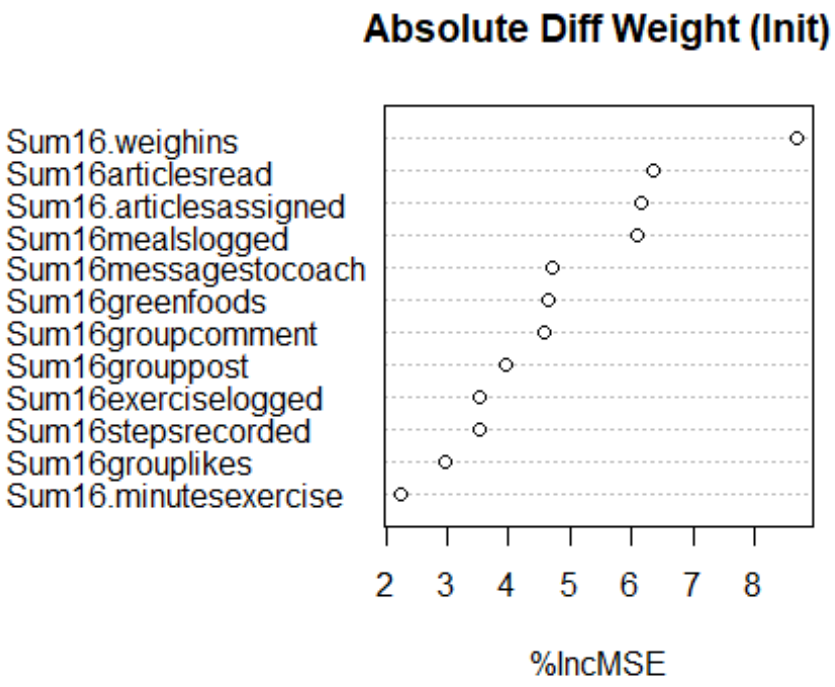
```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
```

```
## Model signifiant at p = 0.01
##    Model R-square:  -0.1353181
##    Random R-square:  -0.2803115
##    Random R-square variance:  0.0004745628
```

Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "Absolute Diff Weight (Init)")
```



Multiple Regression 4

Curriculum Week (or Length of Time with DF?)

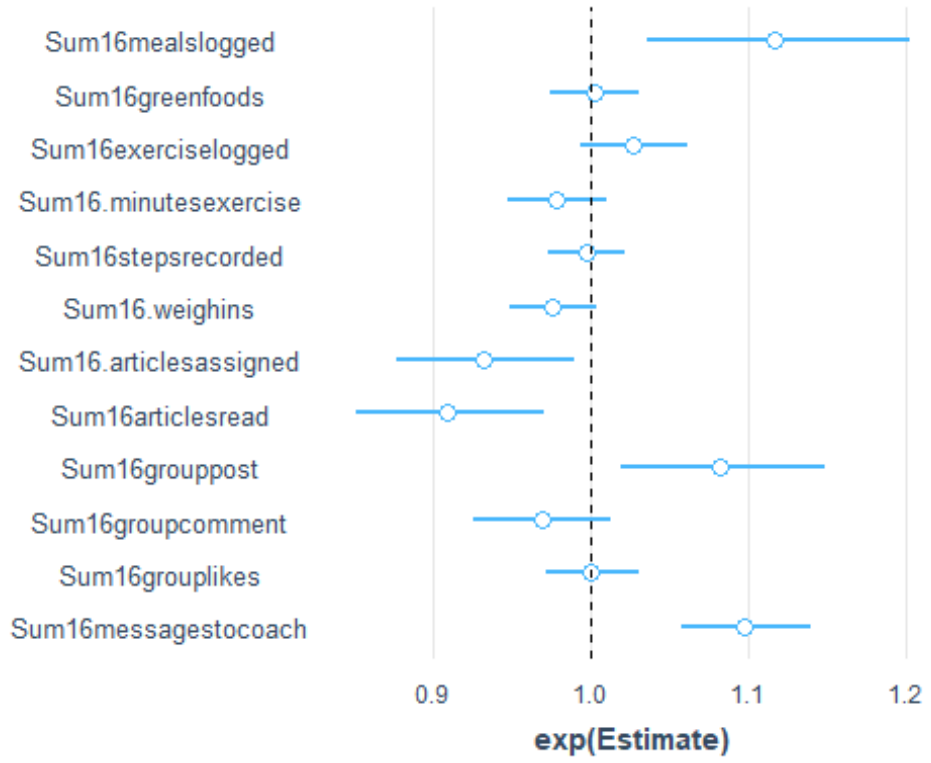
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(17,18,21,24,27,30,33,36,39,42,45,48,51)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(CurriculumWeek ~., DF.t)
```

MODEL INFO:
Observations: 7138
Dependent Variable: CurriculumWeek
Type: OLS linear regression

MODEL FIT:
$F(12,7125) = 7.47$, $p = 0.00$
$R^2 = 0.01$
Adj. $R^2 = 0.01$

Standard errors: OLS

	Est.	2.5%	97.5%	t val.	p	partial.r	part.r
(Intercept)	-0.00	-0.02	0.02	-0.00	1.00		
Sum16mealslogged	0.11	0.05	0.17	3.40	0.00	0.04	0.04
Sum16greenfoods	0.00	-0.03	0.03	0.11	0.91	0.00	0.00
Sum16exerciselogged	0.03	-0.01	0.06	1.53	0.13	0.02	0.02
Sum16.minutesexercise	-0.02	-0.05	0.01	-1.44	0.15	-0.02	-0.02
Sum16stepsrecorded	-0.00	-0.03	0.03	-0.21	0.84	-0.00	-0.00
Sum16.weighins	-0.03	-0.05	0.00	-1.76	0.08	-0.02	-0.02
Sum16.articlesassigned	-0.07	-0.13	-0.01	-2.44	0.01	-0.03	-0.03
Sum16articlesread	-0.10	-0.15	-0.04	-3.41	0.00	-0.04	-0.04
Sum16grouppost	0.08	0.04	0.11	4.52	0.00	0.05	0.05
Sum16groupcomment	-0.03	-0.07	0.01	-1.63	0.10	-0.02	-0.02
Sum16grouplikes	0.00	-0.03	0.03	0.00	1.00	0.00	0.00
Sum16messagestocoach	0.09	0.06	0.12	5.81	0.00	0.07	0.07



Regression Tree 4

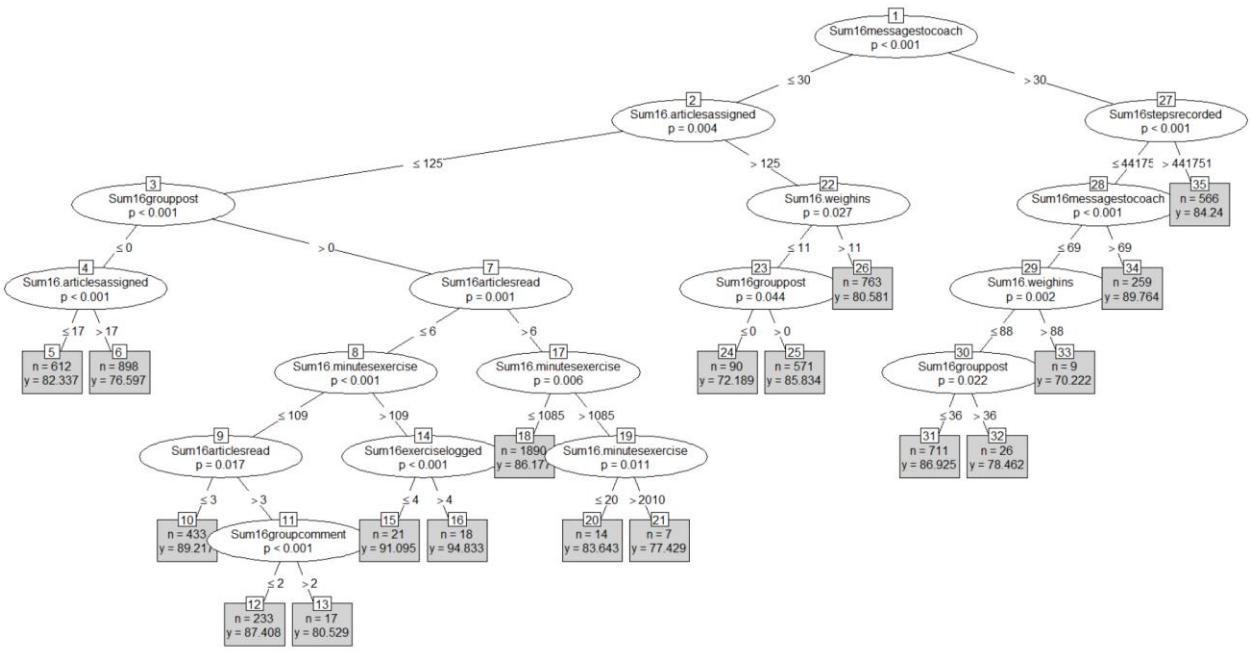
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header = TRUE)
DF <- na.omit(DF)
```

```
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(17,18,21,24,27,30,33,36,39,42,45,48,51)]
model <- train(
  CurriculumWeek ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results
```

```
##   mincriterion     RMSE   Rsquared      MAE    RMSESD RsquaredSD     MAESD
## 1         0.95 14.36489 0.05673412 8.537819 0.5111853 0.02149254 0.2195149
```

Tree Model

```
plot(model$finalModel, type = "simple")
```



User Engagement and Curriculum Week based on 18 Terminal Nodes (Left to Right)

1	Curriculum Week avg 82.34 = **WK 16 Messages to Coach ≤ 30, WK 16 Articles Assigned ≤ 125, WK 16Group Posts ≤ 0, WK 16 Articles Assigned ≤ 17**
2	Curriculum Week avg 76.60 = **WK 16 Messages to Coach ≤ 30, WK 16 Articles Assigned ≤ 125, WK 16 Group Posts ≤ 0, WK 16 Articles Assigned > 17**
3	Curriculum Week avg 89.23 = **WK 16 Messages to Coach ≤ 30, WK 16 Articles Assigned ≤ 125, WK 16 Group Posts > 0, WK 16 Articles Read ≤ 6, WK 16 Minutes Exercise ≤ 109, WK 16 Articles Read ≤ 3**
4	Curriculum Week avg 87.41 = **WK 16 Messages to Coach ≤ 30, WK 16 Articles Assigned ≤ 125, WK 16 Group Posts > 0, WK 16 Articles Read ≤ 6, WK 16 Minutes Exercise ≤ 109, WK 16 Articles Read > 3, WK 16 Group Comment ≤ 2**

5	Curriculum Week avg 80.53 = **WK 16 Messages to Coach ≤ 30, WK 16 Articles Assigned ≤ 125, WK 16 Group Posts > 0, WK 16 Articles Read ≤ 6, WK 16 Minutes Exercise ≤ 109, WK 16 Articles Read > 3, WK 16 Group Comment > 2**
6	Curriculum Week avg 91.10 = **WK 16 Messages to Coach ≤ 30, WK 16 Articles Assigned ≤ 125, WK 16 Group Posts > 0, WK 16 Articles Read ≤ 6, WK 16 Minutes Exercise > 109, WK 16 Exercise Logged ≤ 4**
7 High	Curriculum Week avg 94.83 = **WK 16 Messages to Coach ≤ 30, WK 16 Articles Assigned ≤ 125, WK 16 Group Posts > 0, WK 16 Articles Read ≤ 6, WK 16 Minutes Exercise > 109, WK 16 Exercise Logged > 4**
8	Curriculum Week avg 86.18 = **WK 16 Messages to Coach ≤ 30, WK 16 Articles Assigned ≤ 125, WK 16 Group Posts > 0, WK 16 Articles Read > 6, WK 16 Minutes of Exercise ≤ 1085**
9	Curriculum Week avg 83.64 = **WK 16 Messages to Coach ≤ 30, WK 16 Articles Assigned ≤ 125, WK 16 Group Posts > 0, WK 16 Articles Read > 6, WK 16 Minutes of Exercise > 1085 & ≤ 2010**
10	Curriculum Week avg 77.43 = **WK 16 Messages to Coach ≤ 30, WK 16 Articles Assigned ≤ 125, WK 16 Group Posts > 0, WK 16 Articles Read > 6, WK 16 Minutes of Exercise > 2010**
11	Curriculum Week avg 72.19 = **WK 16 Messages to Coach ≤ 30, WK 16 Step Recorded ≤ 441751, WK 16 Articles Assigned > 125, WK 16 Weigh-ins ≤ 11, WK 16 Group Posts ≤ 0**
12	Curriculum Week avg 85.83 = **WK 16 Messages to Coach ≤ 30, WK 16 Step Recorded ≤ 441751, WK 16 Articles Assigned > 125, WK 16 Weigh-ins ≤ 11, WK 16 Group Posts > 0**
13	Curriculum Week avg 80.58 = **WK 16 Messages to Coach ≤ 30, WK 16 Step Recorded ≤ 441751, WK 16 Articles Assigned > 125, WK 16 Weigh-ins >11**
14	Curriculum Week avg 89.93 = **WK 16 Messages to Coach > 30, WK 16 Step Recorded ≤ 441751, WK 16 Messages to Coach ≤ 69, WK 16 Weigh-ins ≤ 88, WK 16 Group Posts ≤ 36**
15	Curriculum Week avg 78.46 = **WK 16 Messages to Coach > 30, WK 16 Step Recorded ≤ 441751, WK 16 Messages to Coach ≤ 69, WK 16 Weigh-ins ≤ 88, WK 16 Group Posts > 36**
16 Low	Curriculum Week avg 70.22 = **WK 16 Messages to Coach > 30, WK 16 Step Recorded ≤ 441751, WK 16 Messages to Coach ≤ 69, WK 16 Weigh-ins > 88**
17	Curriculum Week avg 89.76 = **WK 16 Messages to Coach > 30, WK 16 Step Recorded ≤ 441751, WK 16 Messages to Coach > 69**
18	Curriculum Week avg 84.24 = **WK 16 Messages to Coach > 30, WK 16 Step Recorded > 441751**

Random Forest 4

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(17,18,21,24,27,30,33,36,39,42,45,48,51)]
rf <- randomForest(CurriculumWeek ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```

Signicance Testing

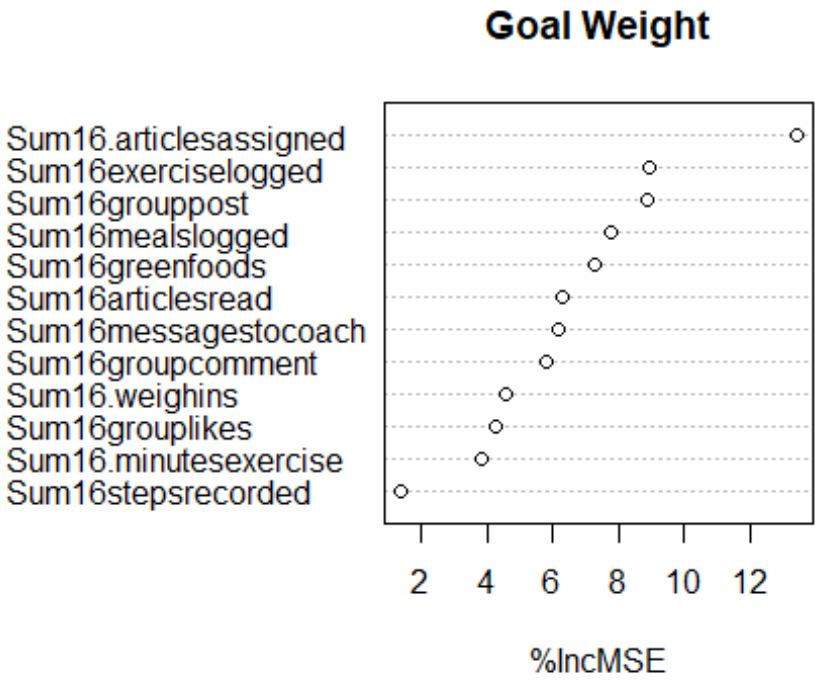
```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##   Model R-square:  -0.1014196
```

```
##    Random R-square:  -0.2758485
##    Random R-square variance:  0.0002591242
```

Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "Goal Weight")
```



Multiple Regression 5

Differnce in BMI

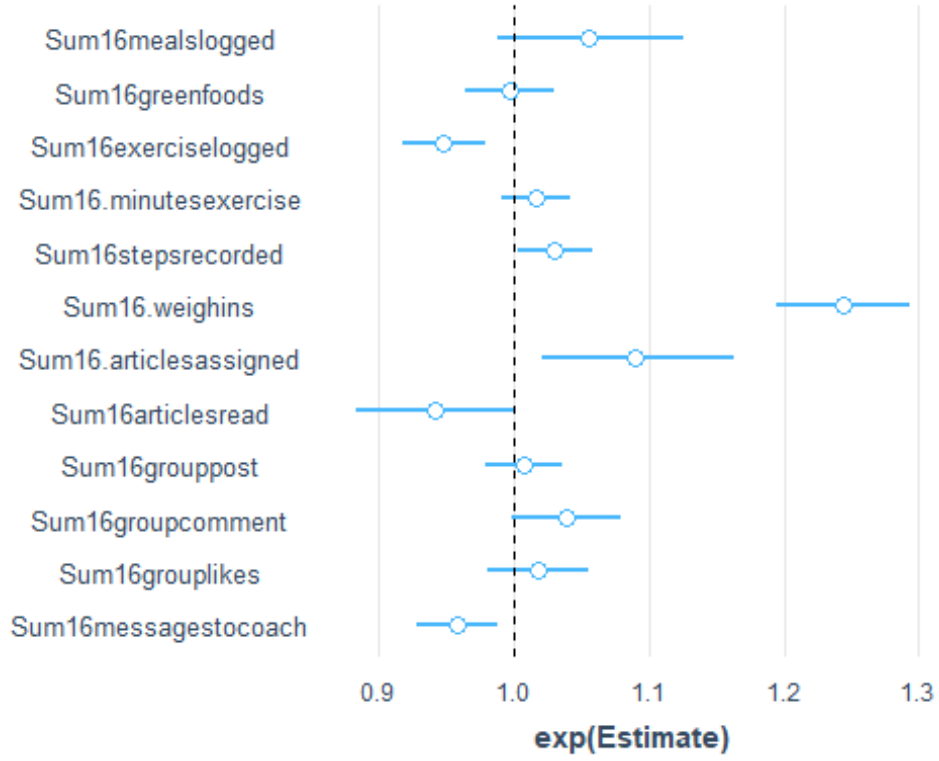
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(14,18,21,24,27,30,33,36,39,42,45,48,51)]
DF.t = scale(DF.t, center = TRUE, scale = TRUE)
DF.t <- as.data.frame(DF.t)
reg <- lm(BMIDifference ~., DF.t)
```

MODEL INFO:
Observations: 7138
Dependent Variable: BMIDifference
Type: OLS linear regression

MODEL FIT:
$F(12,7125) = 48.94$, $p = 0.00$
$R^2 = 0.08$
Adj. $R^2 = 0.07$

Standard errors: OLS

	Est.	2.5%	97.5%	t val.	p	partial.r	part.r
(Intercept)	0.00	-0.02	0.02	0.00	1.00		
Sum16mealslogged	0.05	-0.01	0.12	1.73	0.08	0.02	0.02
Sum16greenfoods	-0.00	-0.03	0.03	-0.22	0.83	-0.00	-0.00
Sum16exerciselogged	-0.05	-0.09	-0.02	-3.21	0.00	-0.04	-0.04
Sum16.minutesexercise	0.02	-0.01	0.05	1.08	0.28	0.01	0.01
Sum16stepsrecorded	0.03	0.00	0.06	2.11	0.03	0.02	0.02
Sum16.weighins	0.22	0.19	0.25	15.73	0.00	0.18	0.18
Sum16.articlesassigned	0.09	0.03	0.14	3.06	0.00	0.04	0.03
Sum16articlesread	-0.06	-0.11	-0.01	-2.21	0.03	-0.03	-0.03
Sum16grouppost	0.01	-0.03	0.04	0.44	0.66	0.01	0.00
Sum16groupcomment	0.04	0.00	0.08	1.98	0.05	0.02	0.02
Sum16grouplikes	0.02	-0.02	0.05	1.02	0.31	0.01	0.01
Sum16messagestocoach	-0.04	-0.07	-0.01	-2.76	0.01	-0.03	-0.03



Regression Tree 5

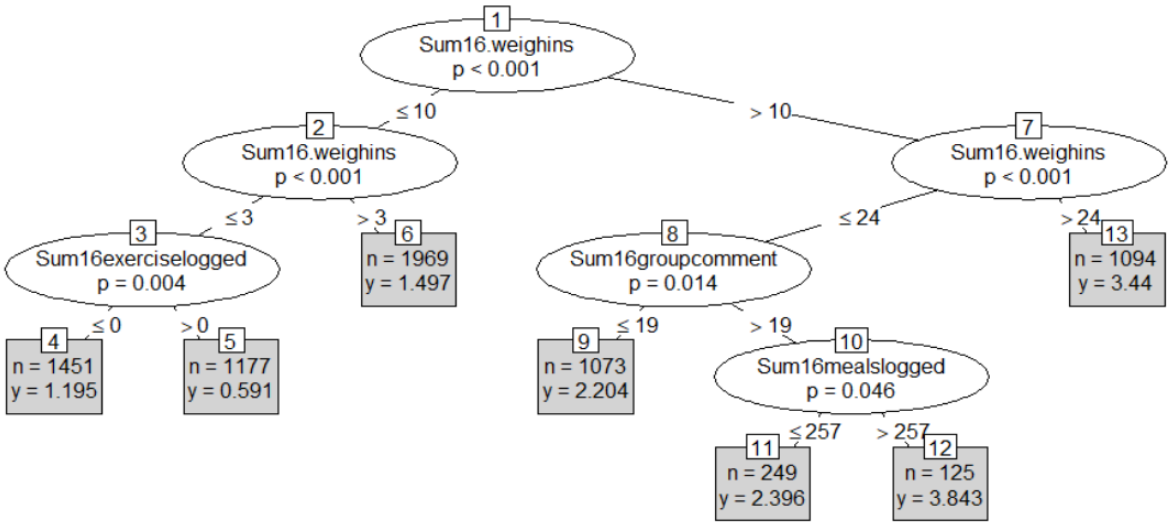
```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
```

```
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(14,18,21,24,27,30,33,36,39,42,45,48,51)]
model <- train(
  BMIDifference ~., DF.t, method = "ctree",
  trControl = trainControl("cv", number = 10),
  tuneGrid = expand.grid(mincriterion = 0.95)
)
model$results
```

```
##   mincriterion     RMSE   Rsquared      MAE    RMSESD RsquaredSD
## 1         0.95 2.958285 0.08314824 1.541024 0.2715949 0.02516501
##        MAESD
## 1 0.06738002
```

Tree Model

```
plot(model$finalModel, type = "simple")
```



User Engagement and Change in BMI for 7 Terminal Nodes (Left to Right)

1	Diff BMI avg 1.20 = **WK 16 Weigh-ins ≤ 3, WK 16 Exercise Logged ≤ 0**
2 Low	Diff BMI avg 0.59 = **WK 16 Weigh-ins ≤ 3, WK 16 Exercise Logged > 0**
3	Diff BMI avg 1.50 = **WK 16 Weigh-ins > 3 & ≤ 10**
4	Diff BMI avg 2.20 = **WK 16 Weigh-ins > 10, WK 16 Weigh-ins ≤ 24, WK 16 Group Comments ≤ 19**
5	Diff BMI avg 2.34 = **WK 16 Weigh-ins > 10, WK 16 Weigh-ins ≤ 24, WK 16 Group Comments > 19, WK 16 Meals Logged ≤ 257**
6 High	Diff BMI avg 3.84 = **WK 16 Weigh-ins > 10, WK 16 Weigh-ins ≤ 24, WK 16 Group Comments > 19, WK 16 Meals Logged > 257**
7	Diff BMI avg 3.44 = **WK 16 Weigh-ins > 10, WK 16 Weigh-ins > 24**

Random Forest 5

```
DF <- read.csv("C:/Users/LaoTz/Desktop/DF Articles/WeightLoss.csv", header =
TRUE)
DF <- na.omit(DF)
DF.t <- DF[-c(36,37,56)]
DF.t <- DF.t[c(14,18,21,24,27,30,33,36,39,42,45,48,51)]
rf <- randomForest(BMIDifference ~ ., data = DF.t, ntree = 25,
                   mtry = 4, nodesize = 5, importance = TRUE)
```

Signicance Testing

```
rf.perm <- rf.significance(rf, DF.t, q = 0.99, p = 0.05, nperm=99, ntree=25)
rf.perm

## Number of permutations:  99
## p-value:  0.01
## Model signifiant at p = 0.01
##   Model R-square:  -0.1625598
##   Random R-square:  -0.2774351
##   Random R-square variance:  0.0004302183
```

Variable Importance Plot

```
varImpPlot(rf, type = 1, main = "BMI Diff")
```